Supplementary Material

Dynamic regulation of gonadal transposon control across the lifespan of the naturally short-lived African turquoise killifish

Bryan B. Teefy[1], Ari Adler[1], Alan Xu[1,2], Katelyn Hsu[1,2], Param Priya Singh[3], Bérénice A. Benayoun[1,2,4,5,6,*]

[1] Leonard Davis School of Gerontology, University of Southern California, Los Angeles, CA 90089, USA.

[2] Molecular and Computational Biology Department, USC Dornsife College of Letters, Arts and Sciences, Los Angeles, CA 90089, USA.

[3] Department of Genetics, Stanford University, Stanford, CA, USA

[4] Biochemistry and Molecular Medicine Department, USC Keck School of Medicine, Los Angeles, CA 90089, USA.

[5] USC Norris Comprehensive Cancer Center, Epigenetics and Gene Regulation, Los Angeles, CA 90089, USA.

[6]USC Stem Cell Initiative, Los Angeles, CA 90089, USA.

[*] Corresponding author (berenice.benayoun@usc.edu).

Supplementary Methods

African turquoise killifish husbandry

African turquoise killifish were raised according to gold standard procedures (Dodzian et al. 2018). To decrease risks of aggression, adult fish were single housed in 2.8L tanks on a recirculating aquatics system manufactured by Aquaneering Inc. System water parameters were as follows: temperature: 29°C; pH: 7.3; conductance: 670-750 µS; ammonia and nitrite: 0ppm and Nitrate: ~15 ppm. Adult fish were fed twice per day with Hikari Freeze Dried Bloodworms 4-6 hours apart, and live *Artemia* once per day. Fry were reared in system water incubated at 28°C, then placed on the recirculating system starting at 2 weeks post hatch. Fry were fed live *Artemia* exclusively until 4 weeks post hatch. The fish facility was kept on a light/dark cycle of 13/11 hours (lights on 9am-10pm). The fish were euthanized by immersion in 1.5 g/L of Tricaine MS-222 dissolved in system water followed by decapitation. All fish were euthanized between 2 and 4 pm to minimize circadian effects. All husbandry conditions and experimental procedures were approved by the University of Southern California (USC) IACUC. Animal care and animal experimentation were performed in accordance with IACUC approved protocols for *Nothobranchius furzeri at* the University of Southern California (approved protocols *20879* and *21023*).

Principal Component Analysis [PCA], Hierarchical Clustering and Variance Partition Analysis

The TEtranscripts count matrices for genes and TEs were analyzed together in R 4.1.2 using DESeq2 v1.34.0 (Love et al. 2014). For global analyses of transcriptional aging patterns across testes and ovaries, all samples were analyzed and normalized together, using sex and age as model covariates (**Fig. 1D-F**; **Supplemental Fig. S1B-S**). The DESeq2 VST-normalized expression log2 count matrices were used as input for analyses of sample relationships: (i) Principal Component Analysis (PCA), (ii) hierarchical clustering with bootstrap resampling with 'pvclust', and (iii) analysis of explained variance as a function of age using 'variancePartition'. For hierarchical clustering, we used 'pvclust' 2.2-0 (Suzuki and Shimodaira 2006), using correlation as a distance metric (*i.e.* 'cor'), average linkage and 1000 bootstrap samples. For the analysis of explained variance by age in ovaries and testes, we leveraged a linear mixed model approach optimized for gene expression studies, as implemented in R package 'variancePartition' 1.24.1 (Hoffman and Schadt 2016). Analyses for piRNA expression were performed similarly, using the count matrix obtained as described in Methods.

Fractional TE analysis

To compare fractional TE expression, the TEtranscripts count matrices used for differential gene and TE expression analysis were imported into R. For each biological replicate, assigned counts were classified as genic or TE-derived and the fractional TE representation for each replicate library was calculated as TE read counts divided by total read counts. Fractional counts were grouped by age and sex and tested for significance by the non-parametric Wilcoxon rank sum test between groups within each sex.

piRNA size distribution and sequence logo analysis

piRNA length distributions (**Supplemental Fig. S6A**) were generated by combining piRNA length histogram data by sex and plotting the resultant length distribution in R. piRNA nucleotide distribution for

each sex was computed by converting piRNA reads to FASTA format files and concatenating by sex. Position weight matrices were generated from these FASTA files in R using the package 'Biostrings' 2.58.0 (**Fig. 5B**). Sequence logos were generated from position weight matrices using the R package 'ggseqlogo' 0.1.

piRNA cluster prediction

We used Protrac 2.42 (Rosenkranz and Zischler 2012) to identify piRNA clusters. Briefly, piRNAs were collapsed to unique reads, while retaining count numbers using the functions TBr2_collapse.pl and TBr2_duster.pl. These reads were then mapped to the soft-masked reference genome described previously using the function sRNAmapper.pl and clusters were identified by a preponderance of piRNA mapping within a particular window using the function proTRAC_2.4.3.pl. Clusters were identified for each biological replicate before being combined such that only unique piRNA clusters remained using the Protrac function merge.pl. To assess the TE composition of piRNA clusters relative to the genome, the intersectbed command (BEDTools 2.27.1) was used to extract piRNA cluster FASTA sequences from the genome (Quinlan and Hall 2010). Then, RepeatMasker 4.1.2-p1 was run as above on the isolated piRNA cluster FASTA sequences to obtain TE family information, which was then contrasted with the genome-wide data in R.

piRNA ping-pong analysis

To measure ping-pong biogenesis, we only considered piRNAs mapping to consensus TE sequences, since they constitute the most biologically relevant target of gonadal piRNAs. To analyze ping-pong biogenesis at a global level, we used the PPmeter 0.4 program (Jehn et al. 2018). Briefly, PPmeter reports the rate of 10 bp overlaps between piRNA reads, a proxy for ping-pong biogenesis, by generating pseudo-replicates through boostrapping one million piRNA reads (100 bootstraps). Using this approach, piRNA libraries can be directly compared to each other by measuring the number of ping-pong events (*i.e.* 10 bp overlaps) per one million piRNA reads. The resultant metric, ping-pong rate per million bootstrapped reads (ppr-mbr) was calculated for each bootstrap in each piRNA library. Bootstrap values were combined per biological replicate and the median value was taken. Median ppr-mbr values were then grouped by age and sex and tested for statistical difference by age group within each sex by the non-parametric Wilcoxon rank sum test.

For TE-specific level ping-pong measurement, we used custom scripts that calculated Z_{10} scores for each consensus TE sequence. Z_{10} scores, as defined in (Han et al. 2015; Vandewege et al. 2022) is the difference of 10 bp overlap occurrences and the mean of all other overlap occurrences within a 20 bp window divided by the standard deviation of all other overlap occurrences. This is shown in the following equation:

$$Z_{10} = \frac{(10\ bp\ overlaps) - (\mu_{1:9,11:20\ bp}\ overlaps)}{\sigma_{1:9,11:20\ bp}\ overlaps}$$

To generate Z_{10} scores, the distances between piRNA 5' ends that mapped to opposite strands of consensus TE sequences were tabulated into a histogram using a custom bash script. The resultant histograms were imported into R and processed in a custom script that reported the Z_{10} score for each consensus TE sequence. For some TE sequences, coverage was incomplete such that we could not derive

Z_{10}-score information; these sequences were discarded from consideration. To measure the changes in Z_{10} scores with age, TE Z_{10} scores were separated by sex and grouped by age. Group-level Z_{10} scores were tested for significant differences by 1-way ANOVA in R, and only TEs with an adjusted p-value <0.05 were considered significantly regulated with aging in either gonad type. Significant TEs were then classified into clusters using 'degPatterns' from 'DEGreport' 1.30.3 package, as previously. For biological group-level analysis of Z_{10} scores, median Z_{10} values over all detected TE sequences were taken for each biological replicate and plotted as a function of age and sex. Difference in Z_{10} scores per group were tested between age groups within sexes by a non-parametric Wilcoxon rank sum test.

Histological analysis of aging ovarian tissue

Ovaries were dissected from 5-week-old, 10-week-old, and 15-week-old female turquoise killifish (N = 2-3 per biological group). Ovaries were fixed in Bouin's solution (Sigma HT10132) at room temperature for 24 h, then washed off using 70% ethanol. Fixed tissue was processed with the Translational Pathology Core at the University of Southern California Norris Comprehensive Care Center for paraffin embedding, sectioning and histological stains. Samples were sectioned longitudinally and for each sample, slides were prepared and stained with hematoxylin and eosin (H&E). Each slide contained 2 tissue sections per slide for technical robustness. Samples were imaged on a Keyence BZ-X710 microscope at 4X magnification.

One section/microphotograph per animal was used to characterize oocyte diameter by four blinded observers using ImageJ/FIJI 2.9.0/1.53t. The distribution of Oocyte diameters at each age group, for each independent observer, is reported in **Supplemental Fig. S4B**, as well as comparison by consecutive age groups (using a goodness-of-fit Kolmogorov-Smirnov test to compare empirical distributions), suggesting similar distribution of oocyte maturation stages irrespective of age.

Although this is a small cohort of fish (N = 8), we believe that large differences in oocyte maturation patterns (like those required to observe changes at the transcriptional level of bulk ovaries) should still be detectable even with these small numbers. Quantified histology pictures have been submitted to Figshare (doi:10.6084/m9.figshare.21572727).

Deconvolution analysis of bulk African turquoise killifish ovarian transcriptomes

To determine whether there were large changes in the cellular composition of killifish ovaries during aging which may underlie transcriptional changes in components of the PIWI pathway, we leveraged a recent single-cell RNA-seq dataset generated on ovaries from young adult zebrafish (Liu et al. 2022). To maximize our ability to obtain accurate deconvolution results, we used the R 'granulator' 1.2.0 pipeline (Pfister et al. 2021), which allows benchmarking and use of 7 state-of-the-art transcriptome deconvolution methods.

To perform the deconvolution process, granulator requires an input matrix of reference gene expression signatures on pure cell types to estimate the relative proportions of each cell type of interest in a target bulk dataset. Thus, we took advantage of the annotations of the single cell RNA-seq zebrafish ovarian dataset that had been deposited to the Broad Institute Single Cell Portal (accession SCP928). Importantly, best deconvolution results are obtained from cell types that are well defined (Pfister et al. 2021), so we collapsed cell types with annotated subsets to their parent cell types (*i.e.* meiotic oocytes, mature oocytes, follicular cells). UMI counts were collapsed by cell type to generate signature expression profiles. In addition,

to test deconvolution accuracy process and benchmark deconvolution methods, we also generated five ovarian *in silico* pseudobulk samples with known proportions of each cell type, randomized from the "true" observed proportions in the zebrafish ovarian single cell dataset.

Granulator requires "transcript per million" (tpm) normalized input RNA-seq data. Thus, for the turquoise killifish input matrix, tpm were obtained by normalizing gene counts from TEtranscripts (see above) relative to transcript lengths, since mRNA-seq covers the entire length of transcripts. For the single cell data, the 10xGenomics protocol is based on sequencing only the sequence upstream poly(A) stretches of transcripts, thus typically covering only ~100bp of the 3'UTR of transcripts. Thus, UMIs were normalized relative to a 100bp length. To "translate" gene expression profiles from the zebrafish context to the turquoise killifish context, we used sequence homology with BLASTP (ncbi-BLAST 2.10.0+) with African turquoise killifish protein sequences from the (Willemsen, 2020) GCA_014300015.1 genome version, against Ensembl release 108 zebrafish protein sequences. For each query zebrafish sequence, only the top hit with E-value $< 10^{-5}$ was retained as its turquoise killifish homolog for deconvolution analysis.

To assess the accuracy of our analyses, we observed that germ cell expression signature profiles clearly cluster apart from that of somatic cell types (**Supplemental Fig. S4C**). Importantly, expression of PIWIL1 is mostly observed in the germline compartment (*i.e.* germline stem and progenitor cells [GSPCs], meiotic oocytes, and maturing oocytes), with maximal expression in more immature germ cells (*i.e.* GSPCs and meiotic oocytes; **Supplemental Fig. S4D**), consistent with the analysis reported in Liu et al., 2022. We benchmarked deconvolution algorithms with 'granulator' to determine whether they were sensitive enough to accurately detect cell types within *in silico* pseudobulk ovarian transcriptional profiles. All benchmarked algorithms performed well (mean Pearson correlation coefficient metric of known *vs.* predicted ovarian cell proportions > 0.88), with support vector regression [SVR] and non-negative least squares [NNLS] as the top 2 performing algorithms (**Supplemental Fig. S4E**).

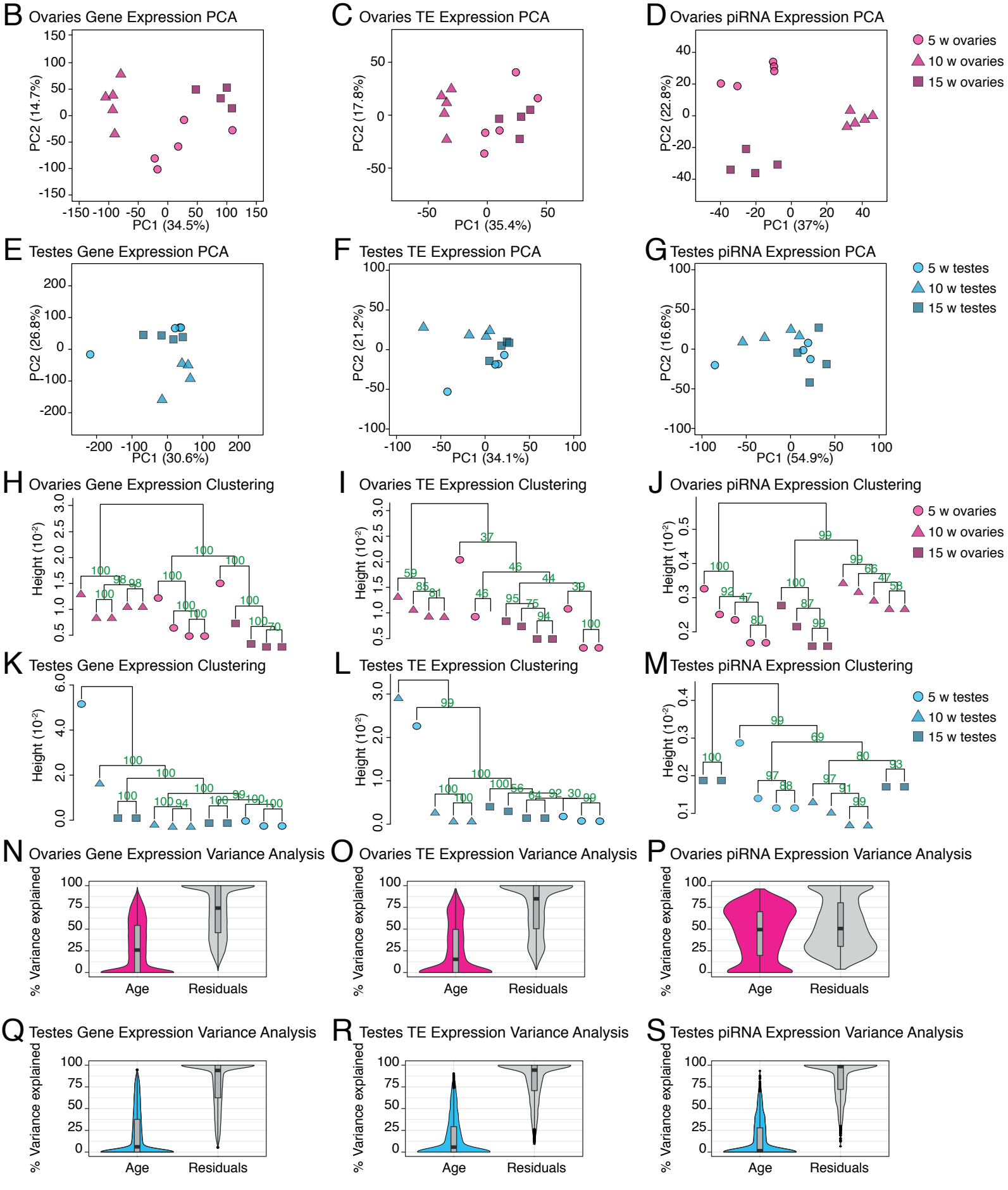
After verifying performance on ovarian *in silico* pseudobulk mixtures, we applied the signature matrix together with SVR and NNLS to the turquoise killifish ovarian transcriptome. Importantly, we did not detect substantial changes in the proportion of immature germ cells between young and middle-aged ovaries by SVR nor NNLS (*i.e.* GSPCs and meiotic cells, which typically show highest expression of PIWI-pathway component genes in young ovaries; **Supplemental Fig. S4D**). Changes in immature germ cells proportion were detected between middle-aged and old ovaries, albeit in opposite directions, and thus not supporting a loss or gain of these cells (**Supplemental Fig. S4F-G**). Thus, our deconvolution analysis suggests that PIWI pathway transcriptional downregulation at middle-age is not driven by broad changes in the ovarian content of immature germ cells.

Supplemental Figures

Figure S1

A

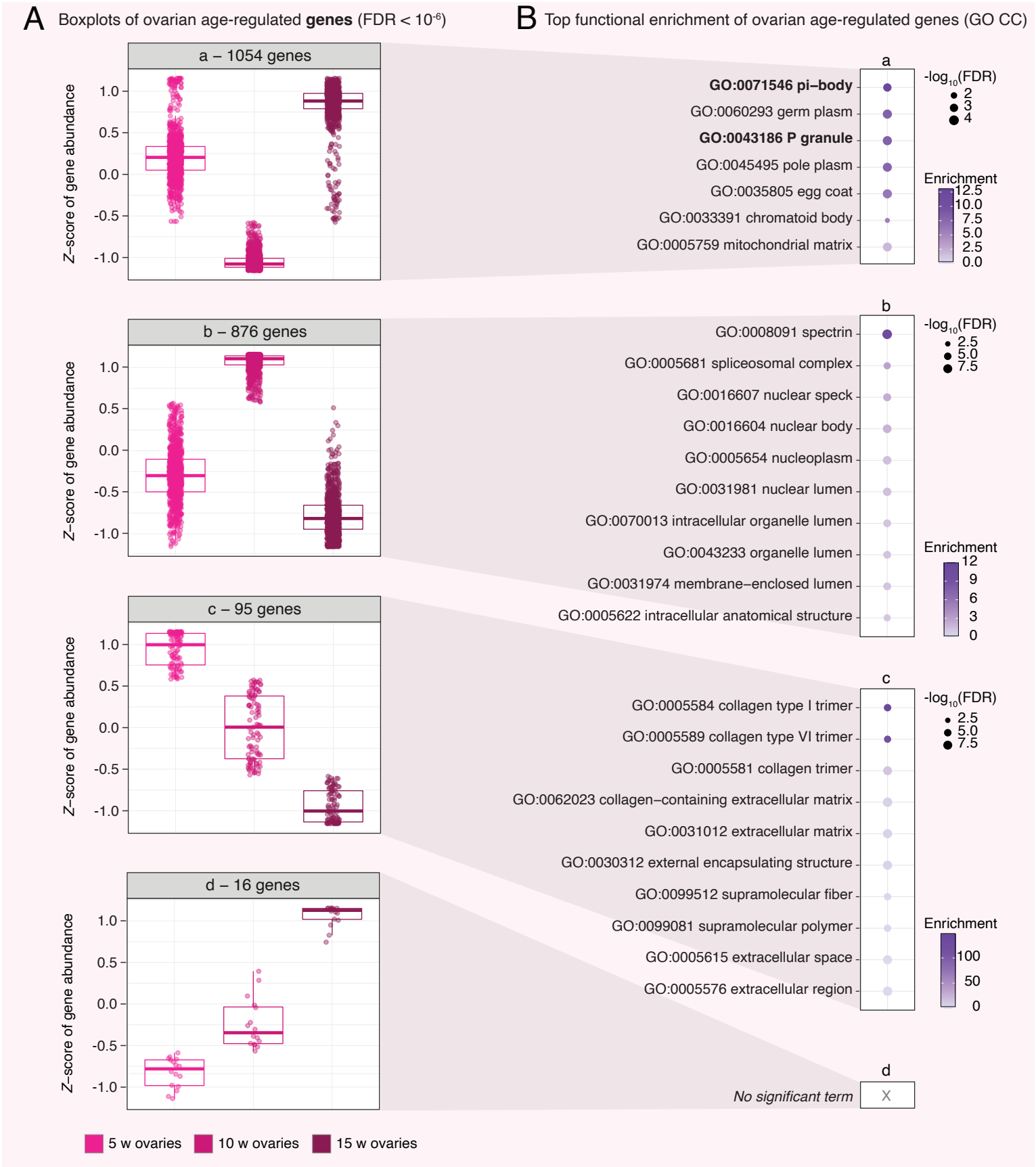
Datasets generated in this study	Ovaries	Testes	Total	
mRNA-seq	5,5,4 (14)	4,4,4 (12)	26	52 datasets
small RNA-seq	5,5,4 (14)	4,4,4 (12)	26	



Supplemental Figure S1. Global characterization of aging turquoise killifish gonadal dataset.

(A) Description of datasets generated in this study. A total of 26 RNA-seq and 26 small RNA-seq datasets were generated from ovary and testes from ages 5, 10, and 15 weeks old. **(B-G)** PCA plots of (B) ovarian gene expression, (C) ovarian TE expression, (D) ovarian TE-targeted piRNA abundance, (E) testicular gene expression, (F) testicular TE expression, (G) testicular TE-targeted piRNA abundance. **(H-M)** pvclust hierarchical clustering of (H) ovarian gene expression, (I) ovarian TE expression, (J) ovarian TE-targeted piRNA abundance, (K) testicular gene expression, (L) testicular TE expression, (M) testicular TE-targeted piRNA abundance. **(N-S)** VariancePartition analysis of variation by age (vs. residual variance) of (N) ovarian gene expression, (O) ovarian TE expression, (P) ovarian TE-targeted piRNA abundance, (Q) testicular gene expression, (R) testicular TE expression, (S) testicular TE-targeted piRNA abundance.

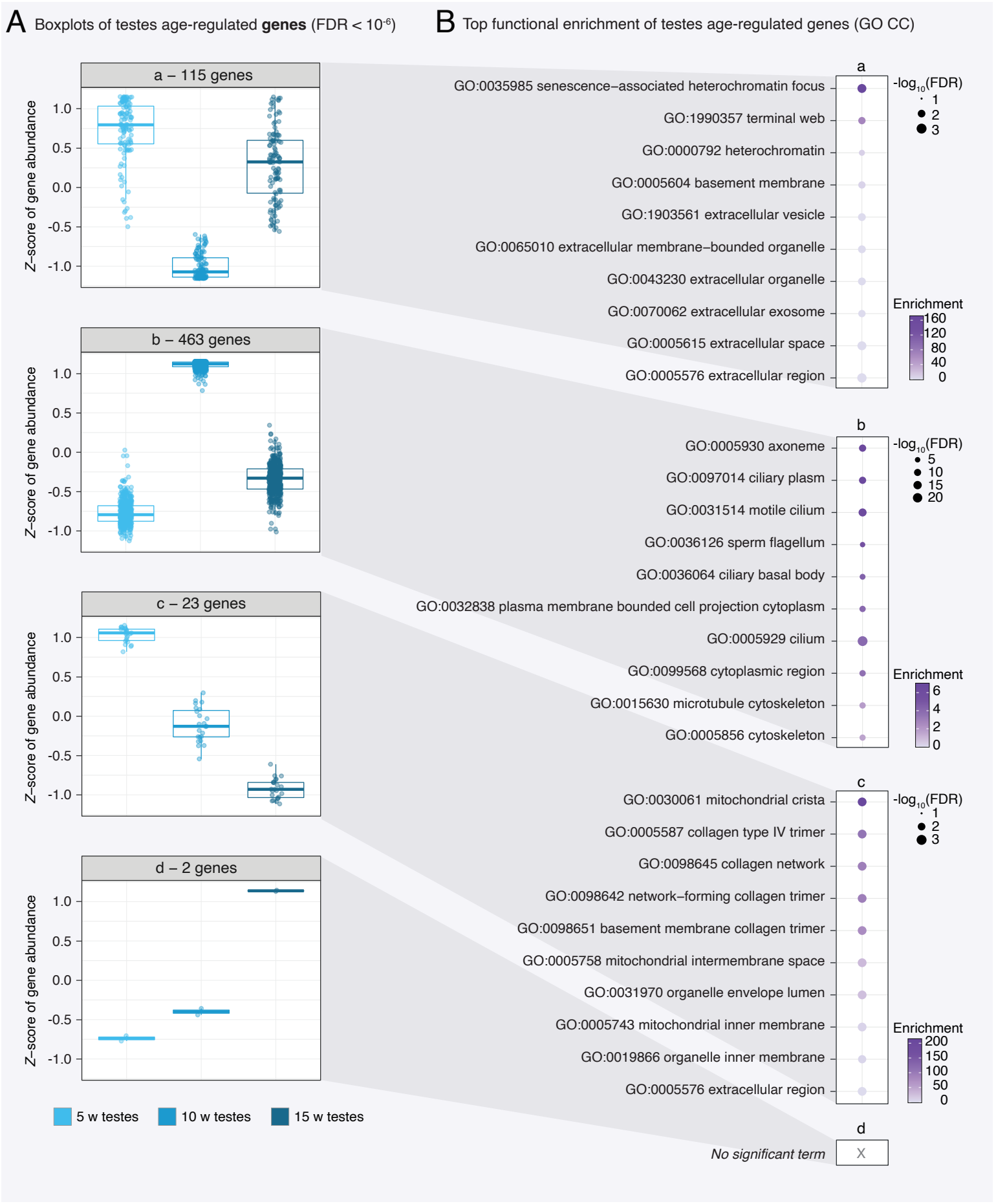
Figure S2


A
Boxplots of ovarian age-regulated genes (FDR < 10-6)
a – 1054 genes
b – 876 genes
c – 95 genes
d – 16 genes
Z–score of gene abundance
1.0
0.5
0.0
-0.5
-1.0
5 w ovaries
10 w ovaries
15 w ovaries
B
Top functional enrichment of ovarian age-regulated genes (GO CC)
a
GO:0071546 pi–body
GO:0060293 germ plasm
GO:0043186 P granule
GO:0045495 pole plasm
GO:0035805 egg coat
GO:0033391 chromatoid body
GO:0005759 mitochondrial matrix
b
GO:0008091 spectrin
GO:0005681 spliceosomal complex
GO:0016607 nuclear speck
GO:0016604 nuclear body
GO:0005654 nucleoplasm
GO:0031981 nuclear lumen
GO:0070013 intracellular organelle lumen
GO:0043233 organelle lumen
GO:0031974 membrane–enclosed lumen
GO:0005622 intracellular anatomical structure
c
GO:0005584 collagen type I trimer
GO:0005589 collagen type VI trimer
GO:0005581 collagen trimer
GO:0062023 collagen–containing extracellular matrix
GO:0031012 extracellular matrix
GO:0030312 external encapsulating structure
GO:0099512 supramolecular fiber
GO:0099081 supramolecular polymer
GO:0005615 extracellular space
GO:0005576 extracellular region
d
No significant term
X
-log10(FDR)
2
3
4
2.5
5.0
7.5
Enrichment
12.5
10.0
7.5
5.0
2.5
0.0
12
9
6
3
0
100
50
0

Supplemental Figure S2. Differential gene expression analysis and GO 'cellular component' enrichment analysis in aging turquoise killifish ovaries.

(A) Boxplots of differentially expressed gene expression groups in ovaries by DESeq2 LRT (FDR < 10^{-6}). Each dot represents the expression level of a significantly differentially expressed gene in each group normalized by *Z*-score for ease of group-to-group comparison. **(B)** Top 10 GO "Cellular Component" terms enriched in each associated cluster according to 'GOstats' (FDR < 5%; see **Supplemental Table S3B** for the complete list of enriched terms). Terms down at middle-age related to the PIWI pathway include "pi-body" and "P granule" terms. Patterns are labelled according to **Fig. 2A**. FDR: False discovery rate. Enrichment: fold enrichment over background.

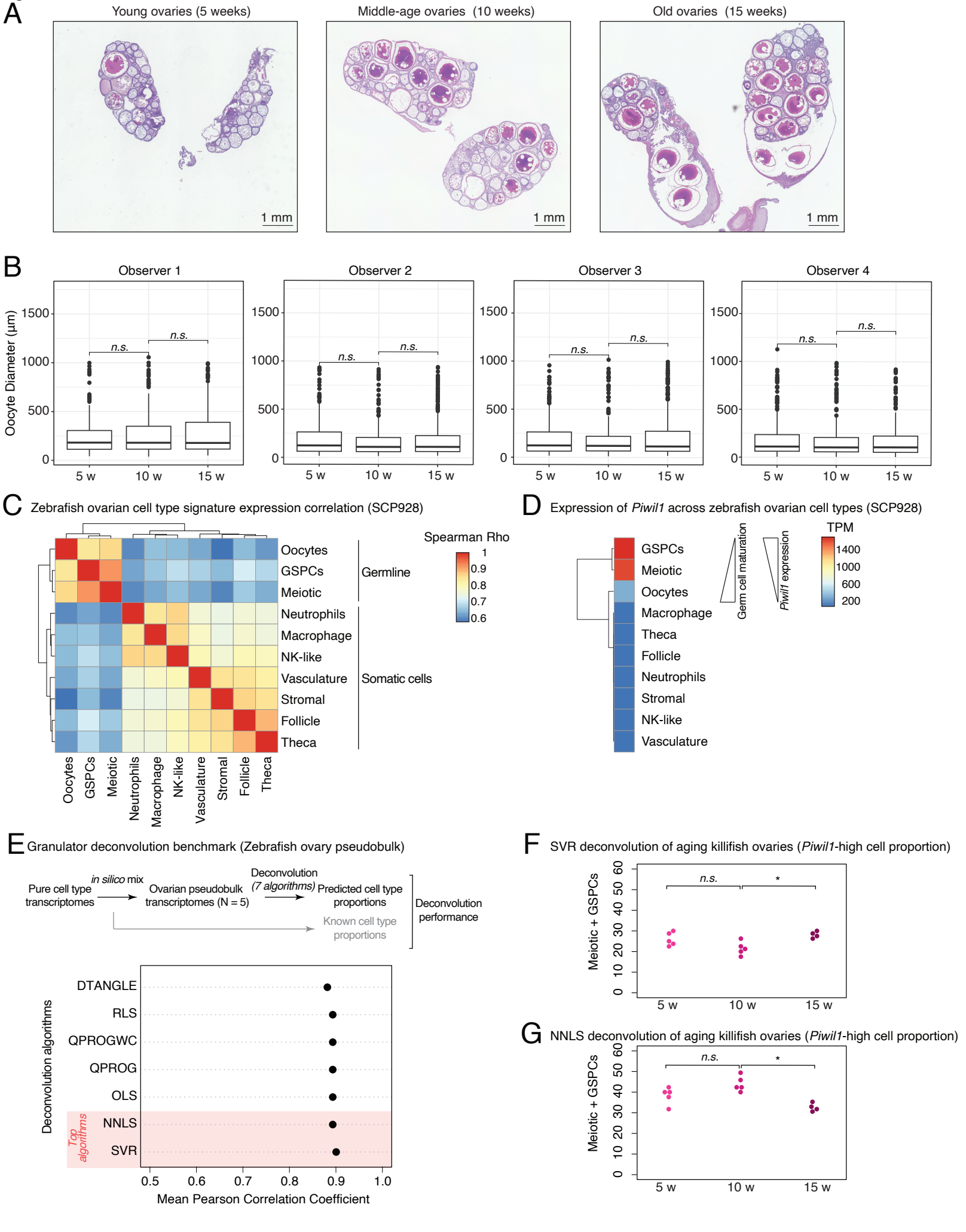
Figure S3


A Boxplots of testes age-regulated genes (FDR < 10-6)
B Top functional enrichment of testes age-regulated genes (GO CC)
a – 115 genes
b – 463 genes
c – 23 genes
d – 2 genes
Z–score of gene abundance
1.0
0.5
0.0
-0.5
-1.0
5 w testes
10 w testes
15 w testes
a
GO:0035985 senescence–associated heterochromatin focus
GO:1990357 terminal web
GO:0000792 heterochromatin
GO:0005604 basement membrane
GO:1903561 extracellular vesicle
GO:0065010 extracellular membrane–bounded organelle
GO:0043230 extracellular organelle
GO:0070062 extracellular exosome
GO:0005615 extracellular space
GO:0005576 extracellular region
-log10(FDR)
1
2
3
Enrichment
160
120
80
40
0
b
GO:0005930 axoneme
GO:0097014 ciliary plasm
GO:0031514 motile cilium
GO:0036126 sperm flagellum
GO:0036064 ciliary basal body
GO:0032838 plasma membrane bounded cell projection cytoplasm
GO:0005929 cilium
GO:0099568 cytoplasmic region
GO:0015630 microtubule cytoskeleton
GO:0005856 cytoskeleton
-log10(FDR)
5
10
15
20
Enrichment
6
4
2
0
c
GO:0030061 mitochondrial crista
GO:0005587 collagen type IV trimer
GO:0098645 collagen network
GO:0098642 network–forming collagen trimer
GO:0098651 basement membrane collagen trimer
GO:0005758 mitochondrial intermembrane space
GO:0031970 organelle envelope lumen
GO:0005743 mitochondrial inner membrane
GO:0019866 organelle inner membrane
GO:0005576 extracellular region
-log10(FDR)
1
2
3
Enrichment
200
150
100
50
0
d
No significant term
X

Supplemental Figure S3. Differential gene expression analysis and GO 'cellular component' enrichment analysis in aging turquoise killifish testes.

(A) Boxplots of differentially expressed gene expression groups in testes by DESeq2 LRT (FDR < 10^{-6}). Each dot represents the expression level of a significantly differentially expressed gene in each group normalized by *Z*-score for ease of group-to-group comparison. **(B)** Top 10 GO "Cellular Component" terms enriched in each associated cluster according to 'GOstats' (FDR < 5%; see **Supplemental Table S3E** for the complete list of enriched terms). Most terms enriched at middle-age are related to spermatogenesis. Patterns are labelled according to **Fig. 2A**. FDR: False discovery rate. Enrichment: fold enrichment over background.

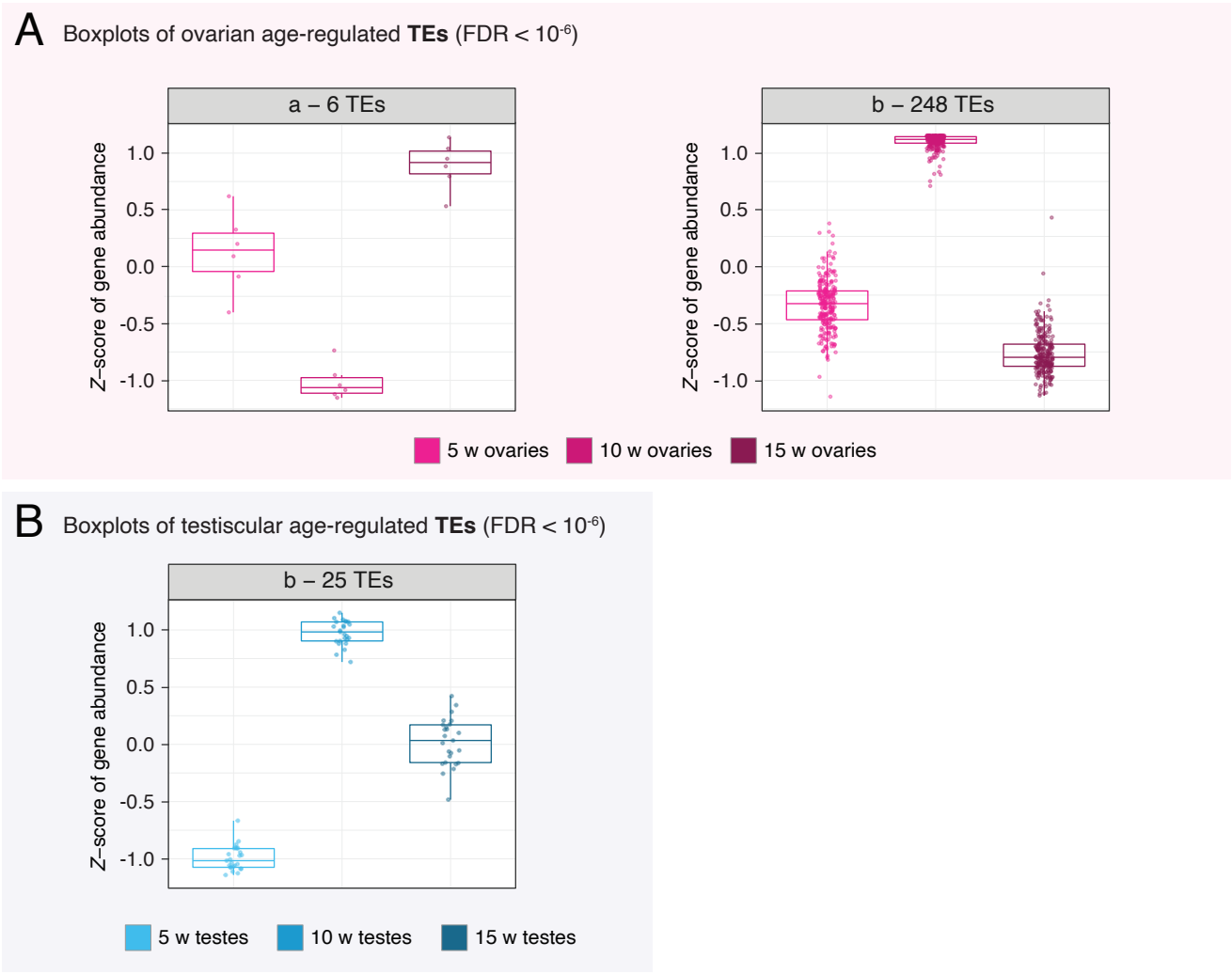
Figure S4


A
Young ovaries (5 weeks)
Middle-age ovaries (10 weeks)
Old ovaries (15 weeks)
1 mm
B
Observer 1
Observer 2
Observer 3
Observer 4
Oocyte Diameter (µm)
n.s.
5 w
10 w
15 w
C
Zebrafish ovarian cell type signature expression correlation (SCP928)
Spearman Rho
Oocytes
GSPCs
Meiotic
Neutrophils
Macrophage
NK-like
Vasculature
Stromal
Follicle
Theca
Germline
Somatic cells
D
Expression of Piwil1 across zebrafish ovarian cell types (SCP928)
TPM
Germ cell maturation
Piwil1 expression
E
Granulator deconvolution benchmark (Zebrafish ovary pseudobulk)
Pure cell type transcriptomes
in silico mix
Ovarian pseudobulk transcriptomes (N = 5)
Deconvolution (7 algorithms)
Predicted cell type proportions
Known cell type proportions
Deconvolution performance
DTANGLE
RLS
QPROGWC
QPROG
OLS
NNLS
SVR
Top algorithms
Deconvolution algorithms
Mean Pearson Correlation Coefficient
F
SVR deconvolution of aging killifish ovaries (Piwil1-high cell proportion)
Meiotic + GSPCs
G
NNLS deconvolution of aging killifish ovaries (Piwil1-high cell proportion)

Supplemental Figure S4. Longitudinal histological and deconvolution analyses of turquoise killifish ovaries.

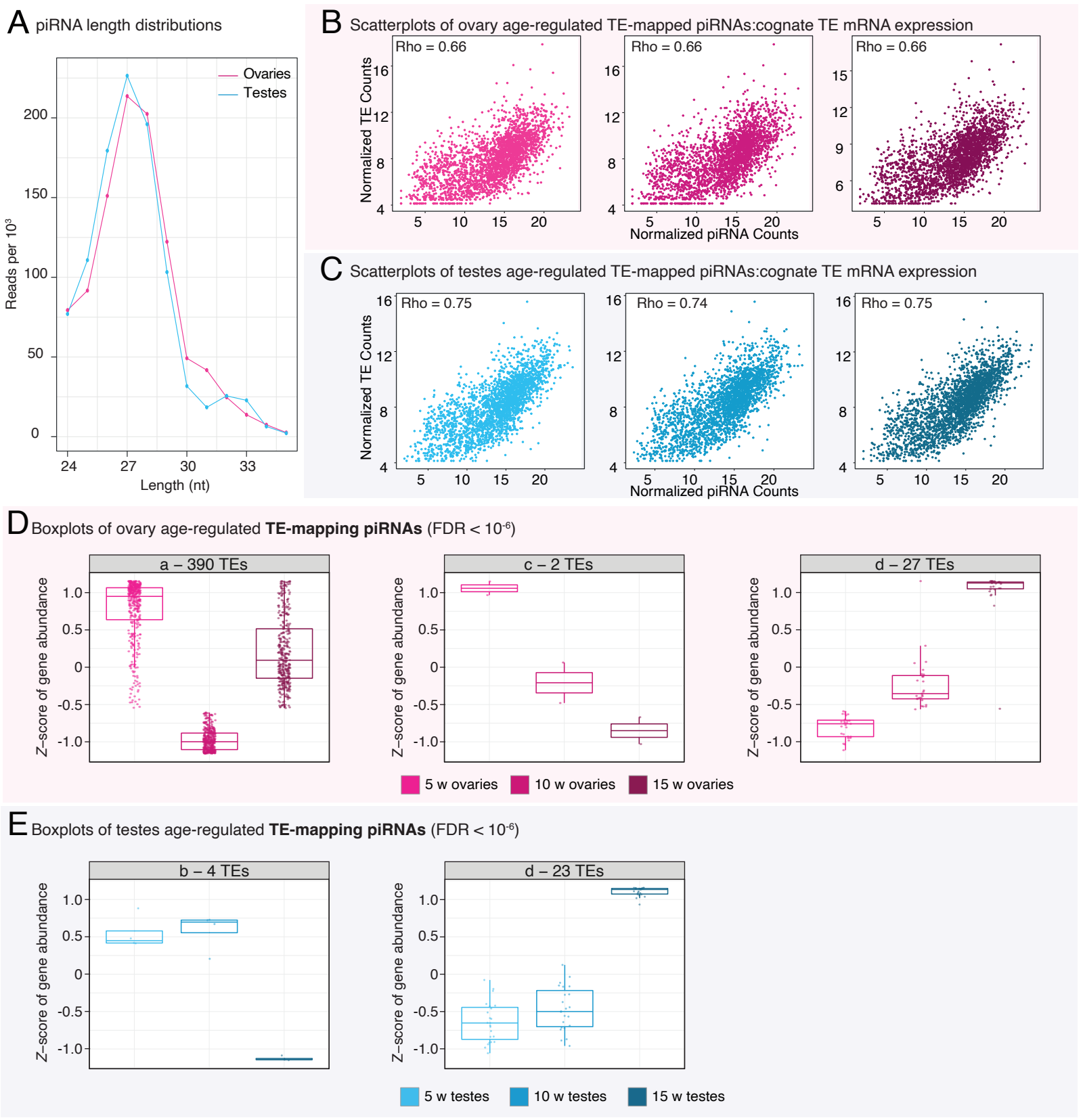
(A) Representative images of 5-, 10-, and 15-week old H&E stained turquoise killifish ovaries. Scale bar: 1mm. **(B)** Boxplots of oocyte diameters grouped by age measured in micrometers by four blinded observers. Distributions of oocyte diameters were compared for young vs. middle-aged ovaries, and middle-aged vs. old ovaries using a Kolmogorov-Smirnov goodness-of-fit test. All comparisons were non-significant (p-value > 0.05), suggesting that the composition of oocytes along maturation stages is largely unchanged throughout turquoise killifish gonadal aging. **(C)** Expression correlation matrix of signature expression profiles from pure ovarian cell types, derived from Zebrafish scRNA-seq dataset from (Liu et al. 2022), using Spearman's rank correlation on TPM expression of killifish homologs. Note the clear separation of germline and somatic cells clearly. **(D)** Expression levels (in tpm) of *Piwil1* in signature expression profiles from pure ovarian cell types show the highest expression in immature germ cells (*i.e.* GSPCs and meiotic oocytes). **(E)** Benchmarking of deconvolution algorithms performance by 'granulator', on *in silico* mixed pseudobulk expression of ovarian gene expression derived from the Zebrafish ovary reference dataset. The mean Pearson correlation coefficient over all cell types for each algorithms are used to reflect the overall performance of each algorithm on the in silico mixes of known proportion. Note that SVR and NNLS show the best performance. **(F-G)** Deconvolution results of the aging African turquoise killifish ovarian bulk transcriptomes using SVR (F) or NNLS (G). The predicted proportion of immature germ cells (*Piwil1*-high) from young, middle-aged, and old deconvoluted turquoise killlifish ovarian bulk transcriptomes. No significant difference was detected between the young and middle-aged samples. Although there were significant changes in the proportions between middle-aged and old ovaries, they are not consistently in the same direction according to both algorithms, Significance in non-parametric Wilcoxon tests. ns: non-significant. *: $p < 0.05$.

Figure S5


A
Boxplots of ovarian age-regulated **TEs** (FDR < 10^{-6})
a – 6 TEs
b – 248 TEs
Z–score of gene abundance
1.0
0.5
0.0
-0.5
-1.0
5 w ovaries
10 w ovaries
15 w ovaries
B
Boxplots of testiscular age-regulated **TEs** (FDR < 10^{-6})
b – 25 TEs
Z–score of gene abundance
1.0
0.5
0.0
-0.5
-1.0
5 w testes
10 w testes
15 w testes

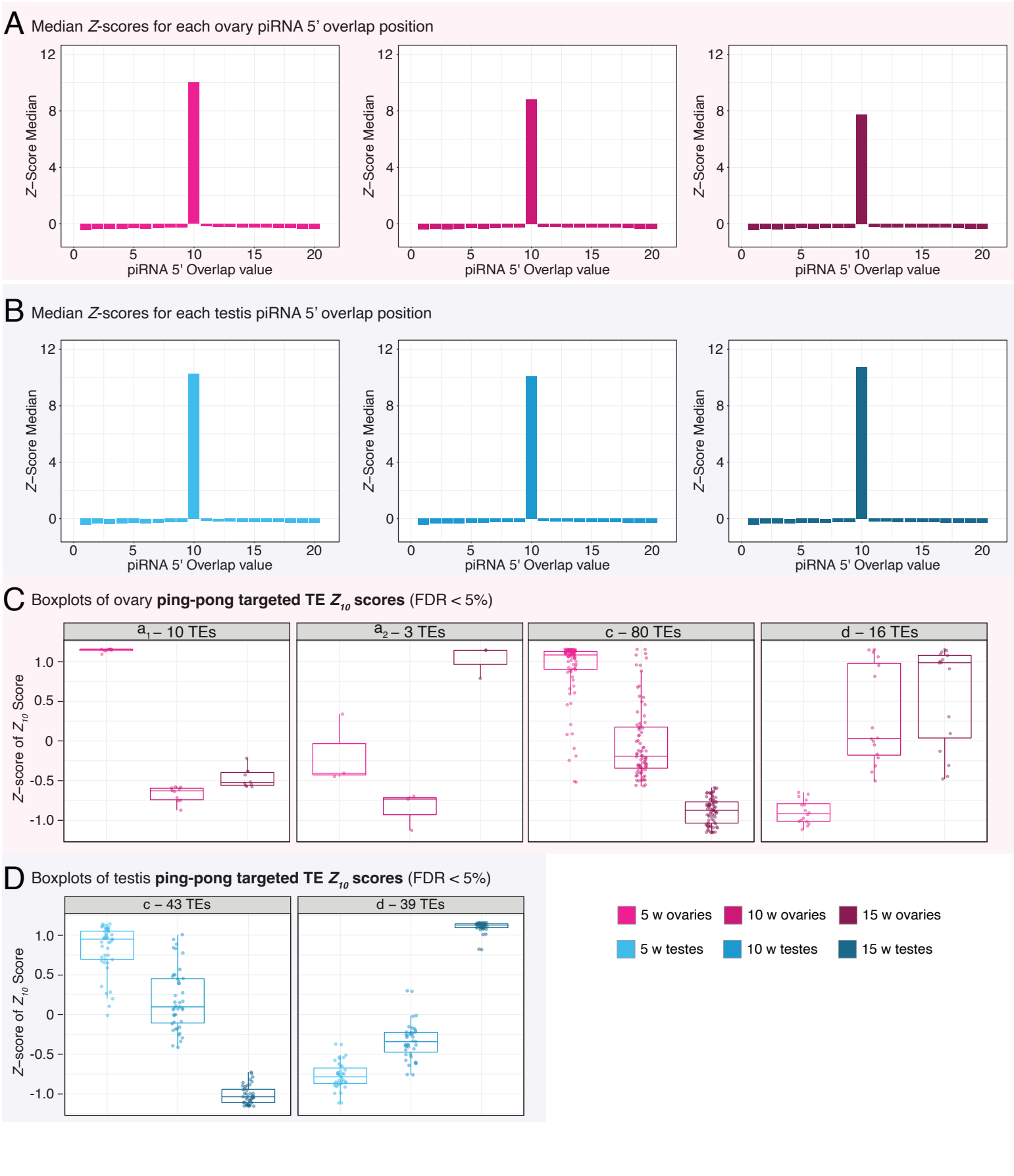
Supplemental Figure S5. TE expression dynamics in the gonads of aging turquoise killifish.
(A) Boxplots of significantly differentially expressed TE clusters in ovaries by DESeq2 LRT (FDR < 10^{-6}). **(B)** Boxplot of the differentially expressed TE clusters in testes by DESeq2 LRT (FDR < 10^{-6}). Each dot represents the expression level of a significantly differentially expressed gene in each group normalized by *Z*-score for ease of group-to-group comparison. Patterns are labelled according to **Fig. 2A**.

Figure S6


A piRNA length distributions
Ovaries
Testes
Reads per 10³
Length (nt)
B Scatterplots of ovary age-regulated TE-mapped piRNAs:cognate TE mRNA expression
Rho = 0.66
Rho = 0.66
Rho = 0.66
Normalized TE Counts
Normalized piRNA Counts
C Scatterplots of testes age-regulated TE-mapped piRNAs:cognate TE mRNA expression
Rho = 0.75
Rho = 0.74
Rho = 0.75
Normalized TE Counts
Normalized piRNA Counts
D Boxplots of ovary age-regulated TE-mapping piRNAs (FDR < 10⁻⁶)
a – 390 TEs
c – 2 TEs
d – 27 TEs
Z–score of gene abundance
5 w ovaries
10 w ovaries
15 w ovaries
E Boxplots of testes age-regulated TE-mapping piRNAs (FDR < 10⁻⁶)
b – 4 TEs
d – 23 TEs
Z–score of gene abundance
5 w testes
10 w testes
15 w testes

Supplemental Figure S6. piRNA characterization and regulation in aging turquoise killifish gonads. **(A)** piRNA length distribution shown by sex, after computational filtering of small RNAs 24-35 bp. Both ovaries and testes have similar piRNA length profiles, with the majority of detected species at ~27bp. **(B)** Scatterplots of young, middle-aged, and old killifish ovaries normalized TE expression *vs.* piRNA counts assigned to the same cognate consensus TE sequence. There is a strong Spearman's rank correlation Rho in all ovarian age groups (Rho = 0.66; significance of correlation test: $p < 2.2x10^{-16}$), indicating that piRNAs are produced sustainably throughout life in response to expressed TEs. **(C)** Scatterplots of young, middle-aged, and old killifish testes normalized TE expression *vs.* piRNA counts assigned to the same cognate consensus TE sequence . There is a slightly higher Spearman's rank correlation in testes than ovaries (Rho = 0.74-0.75; significance of correlation test: $p < 2.2x10^{-16}$), likely reflecting a larger proportion of germ cells in testicular tissue. Rho: Spearman's rank correlation value. **(D)** Boxplots of differentially piRNA-mapped groups of TEs in ovaries, by DESeq2 LRT (FDR < 10^{-6}), using the grouping arrangement defined in **Fig. 2A**. **(E)** Boxplots of differentially piRNA-mapped groups of TEs in testes by DESeq2 LRT (FDR < 10^{-6}).

Figure S7


A Median Z-scores for each ovary piRNA 5' overlap position
B Median Z-scores for each testis piRNA 5' overlap position
Z−Score Median
piRNA 5' Overlap value
C Boxplots of ovary ping-pong targeted TE Z10 scores (FDR < 5%)
a1 – 10 TEs
a2 – 3 TEs
c – 80 TEs
d – 16 TEs
Z−score of Z10 Score
D Boxplots of testis ping-pong targeted TE Z10 scores (FDR < 5%)
c – 43 TEs
d – 39 TEs
5 w ovaries
10 w ovaries
15 w ovaries
5 w testes
10 w testes
15 w testes

Supplemental Figure S7. Analysis of ping-pong biogenesis signatures in aging turquoise killifish gonads (continued).

(A) Median Z_i-scores in ovaries grouped by age measured for each position within a 20 basepair window as summarized for position 10 in **Fig. 6C**. The highest Z_i-score occurs at position 10 (corresponding to the Z_{10} value), indicating that ping-pong is occurring at all ages. **(B)** Median Z_i-scores in testes grouped by age measured for each position within a 20 basepair window. Like the ovarian samples, the highest Z_{10}-score occurs at position 10 indicating ping-pong activity. **(C)** Boxplots of differential Z_{10} scores of TEs grouped into patterns in ovaries (patterns defined in **Fig. 6D**). Each dot is the *Z*-normalized Z_{10} score per age group. **(D)** Boxplots of differential Z_{10} scores of TEs grouped into clusters in testes. Each dot is the *Z*-normalized Z_{10} score per age group.

Supplemental Tables

Supplemental Table S1: Lifespan and fecundity data from GRZ African turquoise killifish reported in this study.

Supplemental Table S2: List of genes with differential expression significance with age by DESeq2 LRT (FDR < 10^{-6}), and expression of "piRNA metabolic pathway" GO term associated genes.

Supplemental Table S3: Enriched GO terms associated to genes with significant age-regulation in ovaries or testes (FDR < 5%).

Supplemental Table S4: List of TEs with differential expression significance with age by DESeq2 LRT (FDR < 10^{-6}).

Supplemental Table S5: List of piRNA targeted TE sequences with differential abundance significance with age by DESeq2 LRT (FDR < 10^{-6}).

Supplemental Table S6: List of piRNA-targeted TE sequences with significant differential Z_{10} ping-pong scores with age by ANOVA (FDR < 0.05).

Supplemental Code

Code used in this study. Also available on GitHub (https://github.com/BenayounLaboratory/Killifish_reproductive_aging_resource).

Supplemental Data

Ovarian histology pictures analyzed in this study. Also available on Figshare (doi:10.6084/m9.figshare.21572727).

Supplementary references

Dodzian J, Kean S, Seidel J, Valenzano DR. 2018. A Protocol for Laboratory Housing of Turquoise Killifish (Nothobranchius furzeri). *J Vis Exp* doi:10.3791/57073.

Han BW, Wang W, Li C, Weng Z, Zamore PD. 2015. Noncoding RNA. piRNA-guided transposon cleavage initiates Zucchini-dependent, phased piRNA production. *Science* **348**: 817-821.

Hoffman GE, Schadt EE. 2016. variancePartition: interpreting drivers of variation in complex gene expression studies. *BMC Bioinformatics* **17**: 483.

Jehn J, Gebert D, Pipilescu F, Stern S, Kiefer JST, Hewel C, Rosenkranz D. 2018. PIWI genes and piRNAs are ubiquitously expressed in mollusks and show patterns of lineage-specific adaptation. *Commun Biol* **1**: 137.

Liu Y, Kassack ME, McFaul ME, Christensen LN, Siebert S, Wyatt SR, Kamei CN, Horst S, Arroyo N, Drummond IA et al. 2022. Single-cell transcriptome reveals insights into the development and function of the zebrafish ovary. *Elife* **11**.

Love MI, Huber W, Anders S. 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* **15**: 550.

Pfister S, Kuettel V, Ferrero E. 2021. granulator: Rapid benchmarking of methods for *in silico* deconvolution of bulk RNA-seq data. R package version 1.2.0. Vol 2022.

Quinlan AR, Hall IM. 2010. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**: 841-842.

Rosenkranz D, Zischler H. 2012. proTRAC--a software for probabilistic piRNA cluster detection, visualization and analysis. *BMC Bioinformatics* **13**: 5.

Suzuki R, Shimodaira H. 2006. Pvclust: an R package for assessing the uncertainty in hierarchical clustering. *Bioinformatics* **22**: 1540-1542.

Vandewege MW, Patt RN, 2nd, Merriman DK, Ray DA, Hoffmann FG. 2022. The PIWI/piRNA response is relaxed in a rodent that lacks mobilizing transposable elements. *RNA* **28**: 609-621.